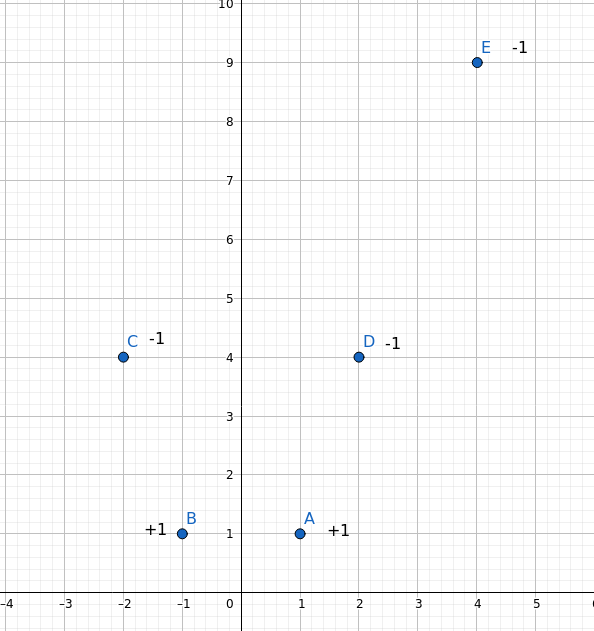
**CS 540: Introduction to Artificial Intelligence**

**HW #4 Solutions to Written Problems**

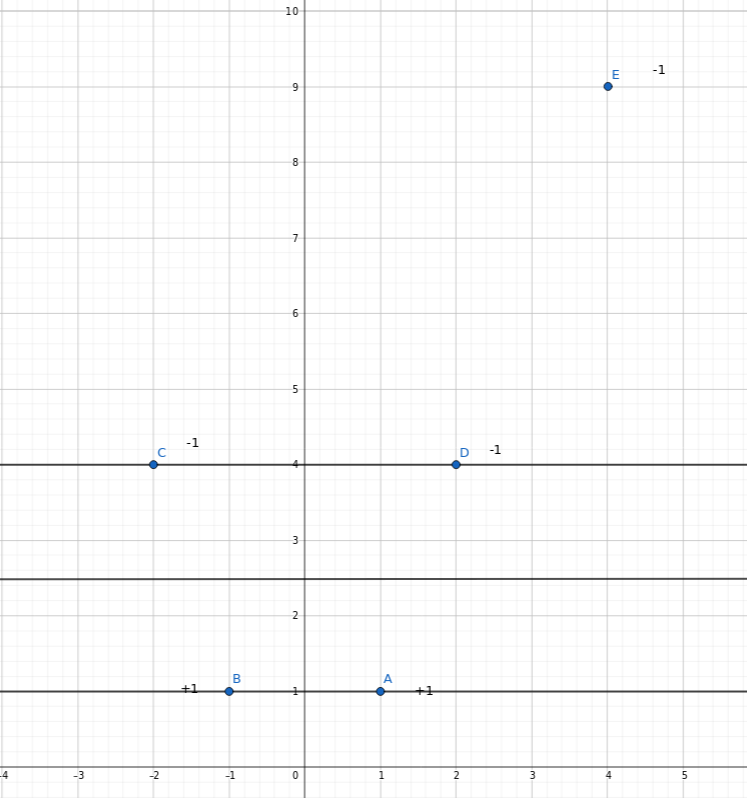
**Problem 1**: Support Vector Machines [10]

1. [5] The transformation of input value X using the mapping function for the given five points results in the following:



Visually from the graph, the support vectors are (-2, 4), (-1, 1), (1, 1) and (2, 4). The line through (1, 1) and (-1, 1) is defined by *q* = 1, or *q* – 1 = 0*.* The line through (-2, 4) and (2, 4) is parallel and defined by *q* = 4, or *q* – 4 = 0. The line midway between these two and parallel to both has the same normal vector **w** = (0, 1) as the other two, and its perpendicular distance from the origin is halfway between 1 and 4 (= 2.5). So, the equation of the decision line is *q* = 2.5, or *q* – 2.5 = 0. So, *w*1 = 0, *w*2 = 1 and *c* = -2.5.

1. [5] The support vectors are the points (-2, 4), (-1, 1), (2, 4) and (1, 1). Given the above equations of the two parallel horizontal lines through these two pairs of support vectors, the normal vector to these two lines is **w** = (0, 1). The margin is defined as the distance from the (center) decision line to the nearest data point, which is the distance from the decision line q = 2.5 to one of the support vectors, say (1, 1), which is **1.5**. If margin is calculated as the distance between the two lines through the support vectors (due to the error in the lecture notes), then the margin is twice that, or **3**. If the margin is computed using the equation 2 /||w||, then the margin is 2 / sqrt(1) = **2**. If the margin is computed using the equation 1 / ||w|| (i.e., the distance between the decision line and one of the lines through support vectors), then the margin is **1**.



**Problem 2:** Neural Networks [20]

1. [5] (A  B)  (A  B)

One solution: w0 = 0.5, wA = 0, wB = -1

The above statement is true when B = false, and is false when B = true. The output is independent of A. There are an infinite number of ways to assign weights, but the weight wB should be a negative number and |wB| should be large enough.

1. [15]

Sigmoid function:

*g*(x) = 1 / (1 + e-*x*)

*g*’(x) = *g*(*x*)(1 – *g*(*x*))

in\_A = (0.3)(0.4) + (0.8)(0.4) + (0.1)(0.4) + 0.3 = 0.78

a\_A = g(0.78) = 0.6857

in\_B = (0.3)(-0.2) + (0.8)(-0.2) + (0.1)(-0.2) + 0.3 = 0.06

a\_B = g(0.06) = 0.515

in\_C = (0.68)(0.4) + (0.515)(-0.2) + 0.3 = 0.4713

a\_C = g(0.4713) = 0.6157

del\_C = (1 – 0.6157)( 0.6157)(1 – 0.6157) = 0.0909

del\_A = (0.6857)(1 - 0.6857)(0.4)(0.0909) = 0.0078

del\_B = (0.515)(1 - 0.515)(-.2)(0.0909) = -0.0045

wac = 0.4 + (0.2)(0.6857)(0.0909) = 0.4124

wbc = (-0.2) + (0.2)(0.515)(0.0909) = -0.1906

wh2c = 0.3 + (0.2)(1)( 0.0909) = 0.3181

wh1a = 0.3 + (0.2)(1)(0.0078) = 0.3015

wh1b = 0.3 + (0.2)(1)(-0.0045) = 0.2991

wx1a = 0.4 + (0.2)(0.3)(0.0078) = 0.4004

wx1b = -0.2 + (0.2)(0.3)(-0.0045) = -0.20027

wx2a = 0.4 + (0.2)(0.8)(0.0078) = 0.4012

wx2b = -0.2 + (0.2)(0.8)(-0.0045) = -0.20072

wx3a = 0.4 + (0.2)(0.1)(0.0078) = 0.4001

wx3b = -0.2 + (0.2)(0.1)(-0.0045) = -0.20009